REMARKS

Applicants have studied the Office Action dated November 17, 2004 and have made amendments to the claims. It is submitted that the application, as amended, is in condition for allowance. By virtue of this amendment, claims 1, 2, 5-12, 14, 17, 18, 20, 21, 25, and 26 are pending. Claims 3, 4, 13, 15, 16, 19, and 22-24 have been canceled without prejudice. Claims 1, 2, 5-12, 14, 17, 18, 20, 21, 25, and 26 have been amended. Reconsideration and allowance of the pending claims in view of the above amendments and the following remarks are respectfully requested.

Claims 1-9 and 12-24 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Brown et al. (U.S. Patent No. 5,913,208). Claims 10, 11, 25, and 26 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Brown et al. in view of "Microsoft Press Computer Dictionary" (Microsoft Press, 1997, p. 309). Claims 3, 4, 13, 15, 16, 19, and 22-24 have been canceled so, with respect to these claims, these rejections are moot. With respect to claims 1, 2, 5-12, 14, 17, 18, 20, 21, 25, and 26, these rejections are respectfully traversed.

The present invention is directed to systems and methods for efficiently comparing and classifying documents. One preferred embodiment of the present invention provides a method for classifying electronically posted documents. According to the method, first and second documents are received, and first and second metadata summaries are generated for the first and second documents. The first metadata summary includes a first plurality of sub-trees and the second metadata summary includes a second plurality of sub-trees, and each of the sub-trees includes a plurality of nodes. The first and second metadata summaries are compared on a structural level by comparing a structure of the sub-trees of the first metadata summary with a structure of the sub-trees of the second metadata summary. If the structures of the sub-trees of the first and second metadata summaries are not equivalent, the first and second documents are identified as distinct. If the structures of the sub-trees of the first and second metadata

summaries are equivalent, a further comparison of the first and second metadata summaries is performed.

The further comparison includes comparing the first and second metadata summaries on a textual level by comparing textual content from the first document that is contained in the sub-trees of the first metadata summary with textual content from the second document that is contained in the sub-trees of the second metadata summary, and identifying the first and second documents as distinct if the textual content from the documents that is contained in the sub-trees of the first and second metadata summaries are not equivalent. Because the first and second metadata summaries are first compared on a structural level, documents with different structures can be quickly identified as distinct without the need to compare attribute values or textual content. Further, because the first and second metadata summaries are compared on a textual level if the structures of the sub-trees are equivalent, documents with the same structure and attributes but with different textual content are correctly identified as distinct.

The Brown reference does not disclose a method in which first and second metadata summaries are generated for first and second documents and a second metadata summary for the second documents with the metadata summaries each including a plurality of sub-trees and each of the sub-trees including a plurality of nodes, the first and second metadata summaries are compared on a structural level by comparing a structure of the sub-trees of the first metadata summary with a structure of the sub-trees of the second metadata summary, the first and second documents are identified as distinct if the structures of the sub-trees of the first and second metadata summaries are not equivalent, and if the structures of the sub-trees of the first and second metadata summaries are equivalent, and a further comparison is performed that includes comparing the first and second metadata summaries on a textual level by comparing textual content from the first document that is contained in the sub-trees of the first metadata summary with textual content from the second document that is contained in the sub-trees of the second metadata summary, as is recited in amended claim 1. Amended claims 12 and 17 contain similar recitations.

Similarly, Brown does not disclose a method in which a first subset of metadata summaries having a first mime-type designation is grouped into a first summary group, first and second metadata summaries are selected from the first summary group, the first and second metadata summaries are compared on a structural level by comparing a structure of the sub-trees of the first metadata summary with a structure of the sub-trees of the second metadata summary, and the first and second documents are identified as distinct if the structures of the sub-trees of the first and second metadata summaries are not equivalent, as is recited in amended claim 10. Amended claim 25 contains similar recitations.

Brown discloses a system in which only attribute values are compared to determine if two documents are equivalent. More specifically, in the system of Brown, an attribute table is made with each row containing the attribute vales for one document, as shown in Figure 3B. Two documents are determined to be identical if certain predefined attributes are the same for the two documents. If the predefined attributes differ, the documents are determined to be distinct.

In contrast, in embodiments of the present invention, a metadata summary is generated for each document, with the metadata summary including a plurality of sub-trees and each of the sub-trees including a plurality of nodes. To determine if two documents are equivalent, their metadata summaries are compared on a structural level by comparing the structures of the sub-trees of their metadata summaries. If the structures of the sub-trees of the metadata summaries are not equivalent, the two documents are identified as distinct. Brown does not teach generating, for each document, a metadata summary that include sub-trees with nodes. Additionally, Brown does not teach comparing the structure of sub-trees generated for two documents to determine if the documents are distinct. Because in embodiments of the present invention the metadata summaries are first compared on a structural level, documents with different structures can be quickly identified as distinct without the need to compare attribute values or textual content. In the system of Brown, document attributes must always be compared.

Further, in the embodiments of the present invention recited in amended claims 1, 12, and 17, if the structures of the sub-trees are equivalent, the metadata summaries are compared on a textual level by comparing textual content from the documents that is contained in the sub-trees of the metadata summaries. The two documents are identified as distinct if the textual content from the documents that is contained in the sub-trees of the metadata summaries are not equivalent. Brown does not teach comparing textual content form documents to determine if the documents are distinct. Brown only ever compares attribute values in determining document equivalency. In fact, Brown teaches away from comparing any actual textual content of the documents. Because in these embodiments of the present invention the metadata summaries are compared on a textual level if the structures of the sub-trees are equivalent, documents with the same structure and attributes but with different textual content are correctly identified as distinct. In the system of Brown, documents with the same attributes but with different textual content are incorrectly identified as the same.

Applicants believe that the differences between Brown, Jacobs, and the present invention are clear in amended claims 1, 10, 12, 17, and 25, which set forth various embodiments of the present invention. Furthermore, the claimed features of the present invention are not realized even if the teachings of "Microsoft Press Computer Dictionary" are incorporated into Brown. "Microsoft Press Computer Dictionary" does not teach or suggest the claimed features of the present invention that are absent from Brown. Therefore, claims 1, 10, 12, 17, and 25 distinguish over the Brown and "Microsoft Press Computer Dictionary" references, and the rejections of these claims under 35 U.S.C. § 103(a) should be withdrawn.

As discussed above, amended claims 1, 10, 12, 17, and 25 distinguish over the Brown and "Microsoft Press Computer Dictionary" references, and thus, claims 2-9, claim 11, claims 13-16, claims 18-24, and claim 26 (which depend from claims 1, 10, 12, 17, and 25, respectively) also distinguish over the Brown and "Microsoft Press Computer Dictionary" references. Therefore, it is respectfully submitted that the rejections of claims 1-26 under 35 U.S.C. § 103(a) should be withdrawn.

In view of the foregoing, it is respectfully submitted that the application and the claims are in condition for allowance. Reexamination and reconsideration of the application, as amended, are requested.

If for any reason the Examiner finds the application other than in condition for allowance, the Examiner is invited to call the undersigned attorney at (561) 989-9811 should the Examiner believe a telephone interview would advance the prosecution of the application.

Respectfully submitted,

Date: April 18, 2005

By:_____

Stephen Bongini
Registration No. 40,917
Attorney for Applicants

FLEIT, KAIN, GIBBONS,
GUTMAN, BONGINI & BIANCO P.L.
One Boca Commerce Center
551 Northwest 77th Street, Suite 111
Boca Raton, Florida 33487
Telephone: (561) 989-9811
Facsimile: (561) 989-9812